[source page in English, except:]

[VOICE] TRANSMISSION DEVICE [AND]
TELEPHONE TO IMPLEMENT IT

19 FRENCH REPUBLIC

NATIONAL INSTITUTE OF INDUSTRIAL PROPERTY

PARIS

11 **Publication No.:** **2 761 800**
(to be used only for reprint orders)

21 **National registration No.:** **97 04521**

51 Int. Cl^6: **G 10 L 9/04**, G 10 L 5/04, 9/20, H 04 M 11/00

12 **INVENTION PATENT APPLICATION** **A1**

22 **Date filed:** 04/02/97

30 **Priority data:**

43 **Application publication date:** 10/09/98
Bulletin 98/41.

56 **List of documents cited in the preliminary search report:** *Please refer to the end of this document.*

60 **References to other related national documents:**

71 **Applicant(s):** *SCANERA SC SOCIETE CIVILE* – FR.

72 **Inventor(s):**

73 **Holder(s):**

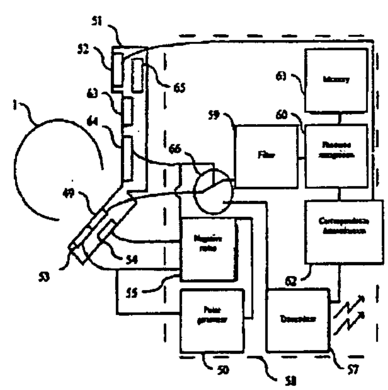
74 **Agent(s):**

54 VOICE TRANSMISSION DEVICE AND TELEPHONE TO IMPLEMENT IT

57 This invention pertains to a voice transmission device that includes a means of detecting the shape of the user's mouth and/or throat (58), by processing whispered sounds, images of the lips or sound echoes.

It also includes a means of determining how each mouth and/or throat shape corresponds to an audible sound (62) and a means of transmitting said audible sound (57).

This invention particularly applies to cellular telephones.



FR 2 761 800 – A1

VOICE TRANSMISSION DEVICE AND TELEPHONE TO IMPLEMENT IT

This invention pertains to a voice transmission device and a telephone to implement it.

Existing voice transmission devices, which use a microphone, are only capable of transmitting the sounds detected by said microphone. They therefore force the user to speak loudly, which is not conducive to discrete voice transmission. This is particularly bothersome in situations involving cellular telephones, which users may wish to use at any time, such as during a meeting or while among people they know or do not know.

This invention is presented as a solution for these disadvantages, as it has a device that is designed to perceive the shape of the vocal organs, specifically the user's throat and/or mouth, and then to associate this shape with an audible sound.

To this end, this invention pertains to a voice transmission device, characterized in that it has a means of detecting the shape of the user's mouth and/or throat, a means of determining how each said mouth and/or throat shape corresponds to an audible sound and a means of transmitting said audible sound.

With these provisions, even when the user only produces sounds of a very low volume, the device perceives the shape of the mouth and/or the throat and associates it with a sound. Indeed, it is a known fact that the shape of the mouth and/or the throat is closely related to the sounds pronounced or articulated by the user.

One specific characteristic is that the means of detecting the shape of the mouth and/or the throat includes a sound receiver adapted to perceive whispers and to transmit a signal that represents said whispers.

With these provisions, the device described in the invention is particularly simple to create, since it uses a sound, ultrasound or infrasound receiver and processes the electric signal transmitted by this sensor.

Another specific characteristic is that the device as briefly described above has filtering means [which are] adapted to prioritize the transmission of frequencies corresponding to whispers.

With these provisions, the device described in the invention extracts the sound frequencies corresponding to the sounds emitted while whispering from among the sounds it receives.

Another specific characteristic is that the means of detecting mouth and/or throat shape includes optical means of sensing mouth shape, adapted to transmit signals that represent said shape, and devices for processing said signals.

With these provisions, the device described in the invention reads the user's lips to determine the sounds he is pronouncing.

Another specific characteristic is that the optical sensor means include a photosensitive sensor matrix.

Thus, by processing images represented by signals coming from the photosensitive sensor matrix, the sounds are read from the user's lips.

Another specific characteristic is that the means of detecting mouth and/or throat shape include an acoustic wave transmitter and an acoustic wave receiver.

With these provisions, the shape detection means function by processing predetermined acoustic wave echoes.

Other specific characteristics include:

- the acoustic wave transmitter is adapted to transmit ultrasound waves, and/or
- the acoustic wave transmitter is adapted to transmit infrasound waves.

With these provisions, the sounds transmitted by said transmitter are not audible and therefore do not disturb the user or his neighbors.

Another specific characteristic is that the device described in the invention as briefly explained above includes two sound sensors, one of which is positioned facing the user's mouth and the other not positioned facing the user's mouth, and a comparison means adapted to transmit, using sound transmission means, signals representing the differences between the signals transmitted by said sensors.

With these provisions, the device is only very slightly disrupted by background sound waves.

The invention also pertains to a telephone, characterized in that it includes a device as briefly described above.

Other advantages, purposes and characteristics of the invention are explained in the description below in reference to the appended drawings, in which:

- Figure 1 depicts a first method of implementing this invention;
- Figure 2 depicts a variation of the first method of implementing this invention;
- Figure 3 depicts a second method of implementing this invention;
- Figure 4 depicts a third method of implementing this invention; and
- Figure 5 depicts an electronic circuit included in each of the methods of implementation presented in reference to Figures 1 through 4.

In the first method of implementing this invention, illustrated in Figure 1, a telephone device (11) includes a keypad (23), a graphic screen (24), a speaker (12), a voice microphone (13) positioned in front of the user's mouth (1), a background microphone (14), a noise reduction circuit (15), a speaker (16) positioned in front of the user's mouth (1), a radio transmitter (17) and a specialized electronic circuit (18). As an explanation, the specialized electronic circuit (18) is depicted in the form of a block illustration, outside of the telephone (11), while in the actual method of implementation it is included inside said telephone (11).

The user's mouth is depicted here in the illustration in the form of a sound cavity.

The specialized electronic circuit (18) has a frequency filter (19) connected to the voice microphone (13), a spectrum recognition circuit (20), a characteristic spectrum memory (21) and a correspondence determination circuit (22).

The telephone device (11) has all the known circuits necessary for the usual functioning of a cellular telephone used with a raised voice, including any electronic agenda, organizer, fax or messaging functions. For this purpose, it has a keypad (23) and a graphic screen (24) that can display graphical symbols, such as alphanumeric symbols, and a telephone data memory (25) that holds telephone numbers corresponding to names.

The speaker (12) is adapted to transmit the sounds received from a telephone interlocutor, according to existing techniques.

The voice microphone (13), positioned in front of the user's mouth (1), is adapted to transmit a signal that represents audible and inaudible frequencies from whispers coming from the user's mouth (1).

The background microphone (14), the noise reduction circuit (15) and the speaker (16) positioned in front of the user's mouth (1) are of the type currently in existence. They use a technique known as "negative noise," which consists of having the speaker (16) emit sound waves in opposite phase with the background sound waves perceived by the background microphone (14), such that in a given area, in this case the user's mouth (1), the sound waves resulting from the combination of the background sound waves and the sound waves emitted by the speaker (16) are of a much lower intensity than that from the background sound waves.

The frequency filter (19), connected to the voice microphone (13), is adapted to reduce the sound frequencies that are not emitted while whispering with a stronger reduction factor than the one applied

to the (audible and inaudible) sound frequencies that are emitted while whispering.

Series of sound spectrums corresponding to the whispering of all phonemes used by the user have been previously stored in the characteristic spectrum memory (21).

The spectrum recognition circuit (20) is adapted to use known methods to extract the spectrum of sound waves coming from the frequency filter and to compare the series of spectrums thus extracted to the series of spectrums stored in the spectrum memory (21). The spectrum recognition circuit (20) transmits a signal that represents the recognized phoneme to the correspondence determination circuit (22).

The correspondence determination circuit (22) transmits sound signals corresponding to:

- the user's voice while pronouncing the phoneme recognized by the recognition circuit (20), and

- the sound intensity, measured as it comes out of the frequency filter (19).

To this end, characteristic elements of the user's voice have been previously stored in the correspondence determination circuit (22), such as the spectrum of sound frequencies that correspond to each of these phonemes, for example, or the sound wave that the user's mouth (1) emits when he pronounces said phoneme.

The signal representing sound waves is transmitted by the correspondence determination circuit, both to the speaker (12), such that the user hears the voice that his interlocutor hears, and to the transmitter (17).

The radio transmitter (17) is part of a radio transmitter/receiver of the type currently found in mobile telephony. It is adapted to remotely transmit electromagnetic waves that represent signals coming from the correspondence determination circuit (22) and to receive electromagnetic waves that represent sound signals transmitted by the interlocutor on the line.

A switch (26), controlled by a button on the keypad (23), switches the output of the voice microphone (13) either as an electrical connection with the frequency filter (19) or directly with the transmitter (17), so the telephone can function as a raised voice telephone.

According to one variation represented in Figure 2, the background sound wave spectrum is compared to that of the waves coming from the user's mouth (1) in order to deduce, on the one hand, which are the whispered sounds, and on the other, the shape of the user's mouth and/or throat. Indeed, aside from the functioning described above in regard to Figure 1 concerning whispering, for each phoneme and for each speaker [user], the mouth and the throat assume a particular shape, and the echoes from the background sounds in this particular shape create a selective reduction of the frequencies present.

To this end, the microphones (13 and 14) are each connected to a spectrum extraction circuit (2 and 3 respectively), and a comparator (4), connected to the outputs of the spectrum extraction circuits, is adapted to normalize the sound intensities and to extract the frequencies that differ most between the two spectrums. The output of this frequency filter is connected to the input of the spectrum recognition circuit (20).

In the second method of implementing this invention illustrated in Figure 3, a cellular telephone (31) includes a speaker (32), a voice microphone (33) positioned facing the user's mouth (1), a keypad (43), a graphic display (44), a telephone data memory (45), a radio transmitter (37), a microcamera (39) positioned facing the user's mouth (1) and a specialized electronic circuit (38).

As an explanation, the specialized electronic circuit (38) is depicted in the form of a block illustration, outside of the telephone (31), while in the actual method of implementation it is included inside said telephone (31).

The specialized electronic circuit (38) has an image processing circuit (47), a phoneme recognition circuit (40), a visual characteristic memory (41) and a correspondence determination circuit (42).

As with existing video cameras, the camera (39) includes a photosensitive sensor matrix.

The image processing circuit (47) is adapted to extract characteristic values from the signal representing the image from the microcamera (39), including the value of the space between the lips in various

vertical and horizontal sections of the mouth. The inventor has shown that by using nine values for sections between the upper part of the upper lip and the lower part of the lower lip, and the width of the lips, the pronunciation of words in French can be differentiated.

The memory (41) has been previously loaded with series of values for said sections and width corresponding to each phoneme pronounced by the user.

The phoneme recognition circuit (40) compares the series stored in the memory corresponding to each phoneme with the series of values measured by the image processing circuit (47). It consequently transmits a digital signal representing the pronounced phoneme to the correspondence determination circuit (42).

Characteristic elements of the user's voice have been previously stored in the correspondence determination circuit (42), such as the spectrum of sound frequencies that correspond to each of these phonemes, for example, or the sound wave that the user's mouth (1) emits when it pronounces said phoneme.

The correspondence determination circuit (42) issues sound signals corresponding to the user's voice when he is pronouncing the phoneme recognized by the recognition circuit (40).

The signal representing sound waves is transmitted by the correspondence determination circuit, on the one hand, to the speaker (32), such that the user hears the voice that his interlocutor hears, and on the other, to the transmitter (37).

The radio transmitter (37) is similar to the radio transmitter (17) depicted in Figure 1. A switch (46), controlled by a button on the keypad (43), switches the input from the transmitter (37) either as an electrical connection with the voice microphone (33) for the usual functioning of existing telephones used with a raised voice, or with the output of the correspondence determination circuit (42).

In another variation that is not depicted, the user's mouth (1) is lit using electroluminescent diodes surrounding the camera lens that emit non-visible light pulses.

In the third method of implementing this invention, depicted in Figure 4, a cellular telephone (51) has an

ear speaker (52), a voice microphone (53) positioned facing the user's mouth (1), a keypad (63), a graphic display (64), a telephone data memory (65), a radio transmitter (57), an ultrasound pulse generator (50), a voice speaker (49) positioned facing the user's mouth (1) and a specialized electronic circuit (58).

As an explanation, the specialized electronic circuit (58) is depicted in the form of a block illustration, outside of the telephone (51), while in the actual method of implementation it is included inside said telephone (51).

The specialized electronic circuit (58) has a frequency filter (59), a phoneme recognition circuit (60), a characteristic memory (61) and a correspondence determination circuit (62).

The pulse generator (50) is adapted to emit pulses at ultrasound frequencies at regular time intervals. The voice speaker (49) is adapted to emit frequencies generated by the pulse generator (50) and those generated by the reduction circuit (55).

The voice microphone (53), positioned in front of the user's mouth (1), is adapted to transmit a signal that represents audible frequencies and frequencies emitted by the voice speaker (49).

The background microphone (54), the noise reduction circuit (55) and the speaker (49) positioned in front of the user's mouth (1) use the technique known as "negative noise," which consists of having the speaker (49) emit sound waves in opposite phase with the background sound waves perceived by the background microphone (54), such that in a given area, in this case the user's mouth (1), the sound waves resulting from the combination of the background sound waves and the sound waves emitted by the speaker (49) are of a much lower intensity than that of the background sound waves, particularly in the frequency range used by the pulse generator (50).

The frequency filter (59), connected to the voice microphone (53), is adapted to reduce the sound frequencies that are not emitted by the voice speaker (49) with a stronger reduction factor than the one applied to the

other (audible and inaudible) sound frequencies that are emitted by the voice speaker (49).

Series of sound spectrums corresponding to the frequencies transmitted by the filter (59) for all phonemes used by the user have been previously stored in the characteristic spectrum memory (61).

The spectrum recognition circuit (60) is adapted to use known methods to extract the spectrum of sound waves coming from the frequency filter and to compare the series of spectrums thus extracted to the series of spectrums stored in the spectrum memory (61). The spectrum recognition circuit (60) transmits a signal that represents the recognized phoneme to the correspondence determination circuit (62).

It is easy to understand that to each shape of the user's mouth and/or throat there corresponds at least one series of sound spectrums coming from the filter (59). This series is, in fact, representative of the absorption of the sound frequencies in the mouth and/or throat cavity and the distances covered by these waves before reaching the voice microphone (53).

The correspondence determination circuit (62) transmits sound signals corresponding to:

- the user's voice when pronouncing the phoneme recognized by the recognition circuit (60), and

- the sound intensity measured at the output of the frequency filter (59).

To this end, characteristic elements of the user's voice have been previously stored in the correspondence determination circuit (62), such as the spectrum of sound frequencies that correspond to each of these phonemes, for example, or the sound wave that the user's mouth (1) emits when it pronounces said phoneme.

The signal representing sound waves is transmitted by the correspondence determination circuit (62), on the one hand, to the ear speaker (52), such that the user hears the voice that his interlocutor hears, and on the other, to the transmitter (57).

The radio transmitter (57) is part of a radio transmitter/receiver of the type currently found in mobile telephony. It is adapted to remotely transmit electromagnetic waves that represent signals coming from the correspondence determination circuit (62) and to receive electromagnetic waves that represent sound signals transmitted by the interlocutor on the line.

A switch (66), controlled by a button on the keypad (23), electrically connects the output of the voice microphone (53) either to the frequency filter (59) or, for functioning at a low voice or with no voice, or directly to the transmitter (57), for the functioning of the raised voice telephone, [sic] The pulse generator (50) is controlled by the same button on the keypad (23) to only emit pulses when the output of the voice microphone (53) is connected to the frequency filter (59).

As a variation, the pulse generator (50) and the voice speaker (49) are jointly adapted to emit infrasound.

In each of the methods of implementation described above, the functioning of the device described in the invention requires prior memory storage of the pronunciation characteristics (frequencies emitted, in the first method of implementation, shape of the lips in the second method of implementation and shape of the mouth and/or throat in the third method of implementation) and the user's voice [characteristics]. However, by default, an initial memory of pronunciation and voice characteristics can be provided that generally apply to all users.

Figure 5 represents an electronic circuit included in each of the methods of implementation described in reference to Figures 1 through 4. This circuit includes the following components, connected to a bus (80): a microprocessor (81), a read-only memory (82) which stores functioning instructions for the microprocessor (81), a random access memory (83) in which the values of variables used while the microprocessor (81) is functioning are stored, such as the measured or detected characteristic values (frequencies emitted, in the first method of implementation, shape of the lips in the second method of implementation and shape of the mouth and/or throat in the third method of implementation), an output port (84) connected to a digital-analog converter (85), which is in turn connected both to a transmitter (86) and to the ear speaker (90)

of the telephone, an input port (87) connected to a rapid processing circuit (88), which is in turn connected to an analog-digital converter (89), which is then in turn connected to a sensor (91) (microphone or camera, depending on the method of implementation), possibly through the intermediary of a frequency filter, as described above.

The rapid processing circuit is adapted to extract the spectrum or process the image, depending on the method of implementation, according to known techniques.

The programming and functioning of the microprocessor (81) and the circuit depicted in Figure 5 do not present any technical difficulties and are therefore not detailed herein.

It should be noted that in each method of implementation, the shape of the user's mouth and/or throat is detected, the correspondence between the detected shape and an audible sound is determined, and said audible sound is transmitted.

CLAIMS

1/ Voice transmission device, characterized in that it includes a means of detecting the shape of the user's mouth and/or throat (20, 40, 60), a means of determining the correspondence between each shape of the mouth and/or throat and an audible sound (22, 42, 62) and a means of transmitting said audible sound (17, 37, 57).

2/ Voice transmission device described in Claim 1, characterized in that the means of detecting the shape of the mouth and/or throat include a sound receiver (13) adapted to perceive whispers and to transmit a signal representing said whispers.

3/ Voice transmission device described in Claim 2, characterized in that it includes filtering means (19) [which are] adapted to prioritize the transmission of frequencies corresponding to whispers.

4/ Voice transmission device described in any of Claims 1 through 3, characterized in that the means of detecting mouth and/or throat shape includes optical means of sensing (39) mouth shape, adapted to transmit signals that represent said shape, and devices for processing said signals (47).

5/ Voice transmission device described in Claim 4, characterized in that the optical sensor means (39) include a photosensitive sensor matrix.

6/ Device described in any of Claims 1 through 5, characterized in that the means of detecting the shape of the mouth and/or throat include an acoustic wave transmitter (53) and an acoustic wave receiver (49).

7/ Device described in Claim 6, characterized in that the acoustic wave transmitter (53) is adapted to transmit ultrasound waves.

8/ Device described in Claim 6, characterized in that the acoustic wave transmitter (53) is adapted to transmit infrasound waves.

9/ Device described in any of Claims 1 through 8, characterized in that it includes two sound sensors (13, 14, 49, 54), one of which is positioned facing the user's mouth (1) and the other not positioned facing the user's mouth, and a comparison means (4) adapted to transmit, using sound transmission means, signals representing the differences between the signals transmitted by said sensors.

10/ Telephone, (11, 31, 51), characterized in that it includes a device described in any of Claims 1 through 9.

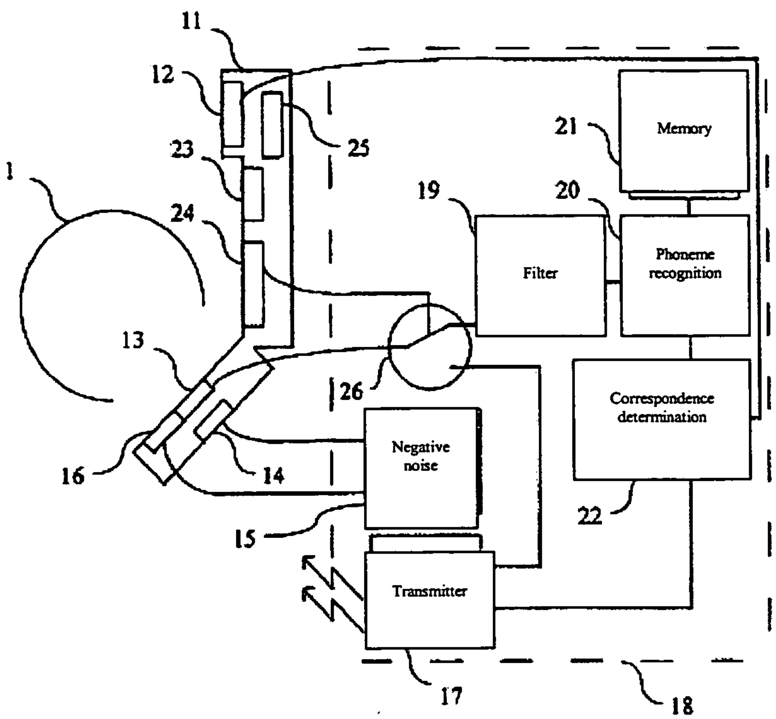

11
12
23
25
1
24
21
Memory
19
20
Filter
Phoneme recognition
13
26
Correspondence determination
Negative noise
16
14
22
15
Transmitter
17
18

Figure 1

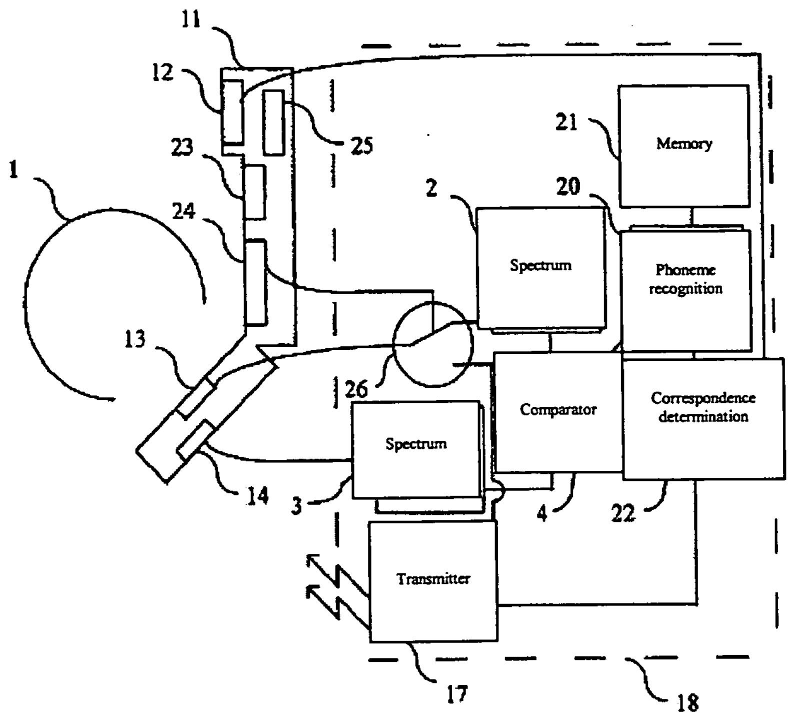

11
12
25
23
21
Memory
1
2
20
24
Spectrum
Phoneme recognition
13
26
Comparator
Correspondence determination
Spectrum
14
3
4
22
Transmitter
17
18

Figure 2

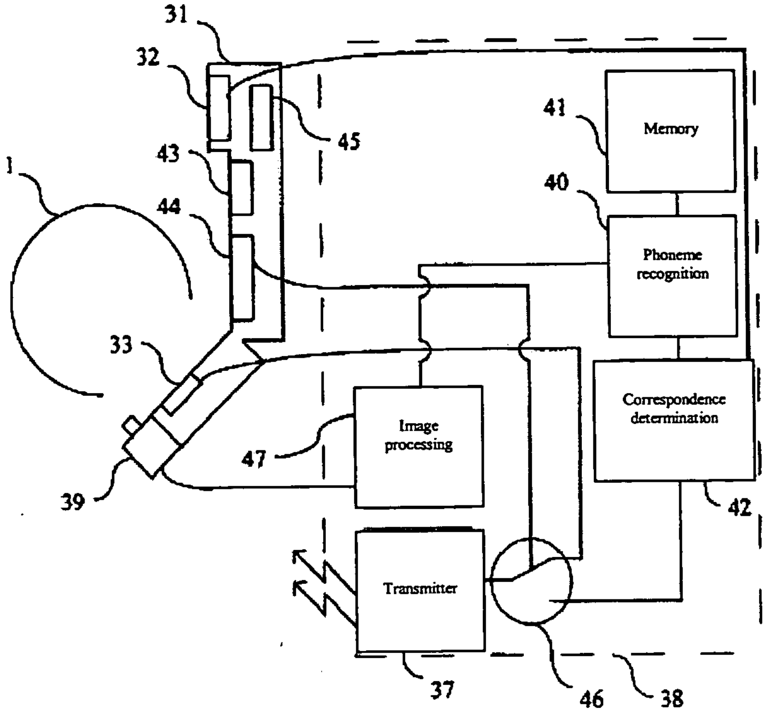

31
32
45
41
Memory
40
43
1
44
Phoneme recognition
33
Image processing
Correspondence determination
47
39
42
Transmitter
37
46
38

Figure 3

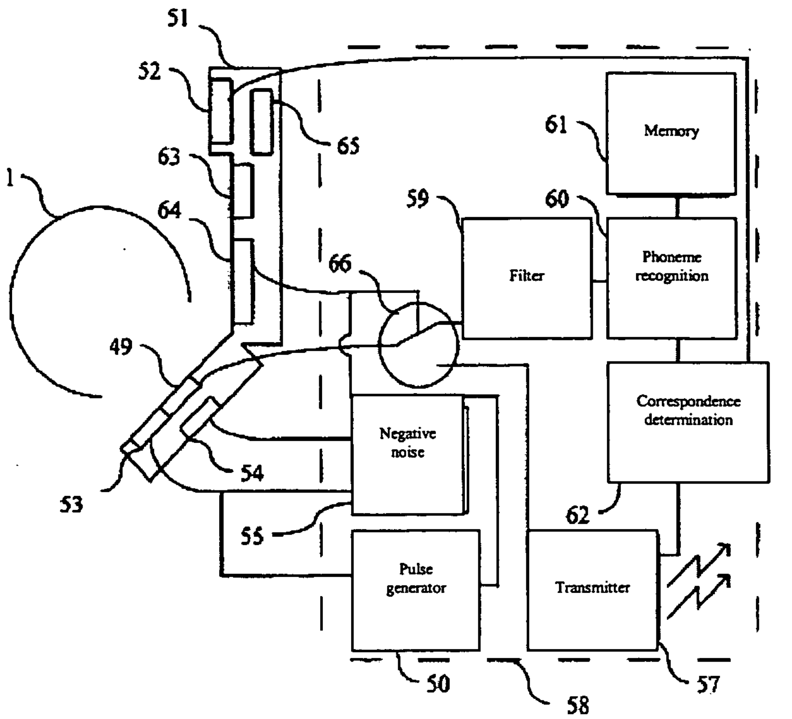

51
52
65
63
1
64
61
Memory
59
60
Filter
Phoneme recognition
66
49
Correspondence determination
Negative noise
54
53
55
62
Pulse generator
Transmitter
50
58
57

Figure 4

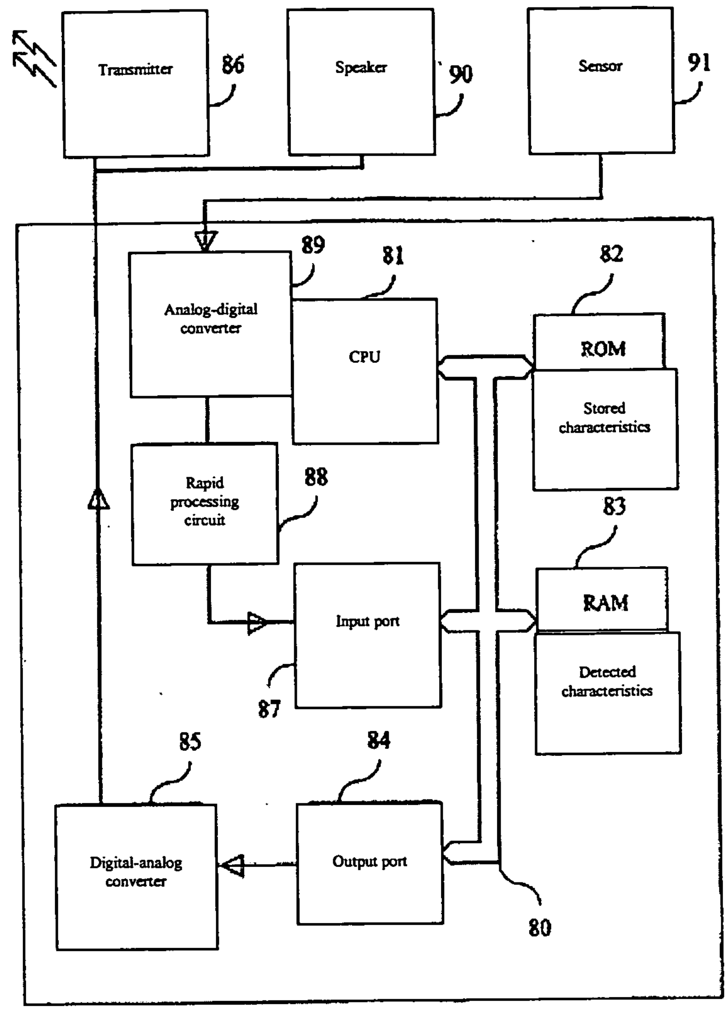

Transmitter
86
Speaker
90
Sensor
91
89
81
82
Analog-digital converter
CPU
ROM
Stored characteristics
Rapid processing circuit
88
83
RAM
Input port
Detected characteristics
87
85
84
Digital-analog converter
Output port
80

Figure 5

FRENCH REPUBLIC

NATIONAL INSTITUTE of INDUSTRIAL PROPERTY

PRELIMINARY SEARCH REPORT

established based on the last claims filed before the search began

National registration No.

FA 543244
FR 9704521

DOCUMENTS CONSIDERED TO BE PERTINENT

Category	Citation of the document with indication, if necessary, of the pertinent parts	Claims pertaining to the application in question
X	"INFRARED NEURO-ANALYZER MICROPHONE (IRNA MICROPHONE)" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 37, no. 8, August 1994, pages 133-135, XP000456190 * page 133 * * page 135, line 14 - line 29; figures 1-3 *	1, 4, 5
Y		9, 10

Y	EP 0 557 166 A (ALCATEL RADIOTELEPHONE) * column 2, line 57 - column 4, line 17; figure 1 *	9, 10

X	MAK M W ET AL: "LIP-MOTION ANALYSIS FOR SPEECH SEGMENTATION IN NOISE" SPEECH COMMUNICATION, vol. 14, no. 3, June 1, 1994, pages 279-296, XP000501478 * page 289, right column, line 1 - page 290, right column, line 2; figure 7 *	1, 4

X	EP 0 332 890 A (IBM) * column 5, line 28 - column 7, line 40; figures 1, 2 *	9

A	EP 0 661 689 A (SONY CORP) * page 2, line 34 - page 3, line 11; figure 1 *	10

TECHNICAL FIELDS RESEARCHED (Int.Cl.6)

G10L

Search completion date	Examiner
December 18, 1997	Pulluard, R.

CATEGORIES OF DOCUMENTS CITED

X: particularly pertinent by itself
Y: particularly pertinent in combination with another document of the same category
A: pertinent with respect to at least one claim or general technological background
O: unwritten disclosure
P: intercalary document

T: theory or principle on which the invention is based
E: patent document with a date prior to the filing date and that was only published as of this filing date or a subsequent date
D: cited in the application
L: cited for other reasons

&: member of the same family, corresponding document

TRANSPERFECT
TRANSLATIONS

Certificate of Accuracy

I, Andrew Green, of TransPerfect Translations Inc., an ISO 9001:2000 Certified company, do hereby declare that the following document, as translated by a team of TransPerfect's linguists, is to the best of my knowledge and belief a true and correct translation of the following "French Patent No. 2761800" from French into English.

I so declare under penalty of perjury on this day, June 2, 2006.

Andrew Green

Andrew Green
Director, Business Development
TransPerfect Translations, Inc.
600 University Street, Suite 1615
Seattle, WA 98101
Phone: (206) 652-4147
Fax: (206) 652-4805